

日本人大学生のモノログ型タスクにおける 英語スピーキング能力の発達の評価

Abstract

With modern technological advances and changes in the social environment, language teachers are responsible for regularly updating their professional knowledge and skills and evaluating their own teaching practice (Richards & Farrell, 2005). This study explores manageable measures that language teachers can use to evaluate their students' development in L2 (English as a second or foreign language) speech. Of the commonly used four indicators of spoken language proficiency (accuracy, fluency, complexity and lexis) (Skehan & Foster, 2012), fluency and lexis are highlighted in this study. Also, drawing on the findings from previous studies on task-based interaction, this study presents a preliminary analysis of 18 university students' L2 speech data elicited in a timed monologue task. The results do not show any clear sign of the students' progress in terms of lexis in three months, but the total number of words produced in the given time slightly increased. The study also discusses insights teachers as researchers would gain through analyzing their students' speech performance.

1. はじめに

近年、言語教育の分野において、教師の職能開発 (professional development) の重要性が認識されるようになってきている (e.g., Richards & Farrell, 2005)。Schön (1984) が提唱した反省的実践 (reflective practice) の概念が広く応用され、特に 2000 年代以降、第二言語および外国語としての英語 (English as a second or foreign language, 以下 L2) の分野でも、教師の自己評価と授業改善を目的とした様々な方法論が生まれている (e.g., Mann & Walsh, 2017)。

語学教師であれば、学期や学年の始めと終わりで、学習者の発話に変化が起きているこ

とをなんとなく感じる場合がある。これを直感でなく、客観的な指標を用いて教師自身が分析・評価する方法はないかという問いが本研究の起点である。具体的には、授業を行いながら、学習者のスピーキング能力の発達を確認できる比較的簡易な査定法を確立し、省察的実践家 (reflective practitioner) として教師が成長するための有効なツールを作り出すことを最終的な目標としている。

L2 話者のスピーキング能力の査定については、応用言語学では、後で述べるように複雑さ (Complexity)、正確さ (Accuracy)、流暢さ (Fluency) という 3 つの指標 (頭文字を取って CAF) が代表的な変数として広く用いられている (Skehan 1998; Ellis 2003, 2009; Ellis and Barkhuizen 2005)。さらに、CAF に加え、語彙 (lexis) という指標を併せて、4 つの指標でスピーキング能力の評価を行うという方法が多くの研究で採用されている (e.g., Skehan & Foster, 2012)。

教師が学習者の発話データを分析して、授業改善に活用できる実践的な方法の確立を目指し、L2 話者のスピーキング能力の評価指標のうち、本稿ではまず「流暢さ」(CAF の「F」)、および「語彙」、特にスピーキングにおける情報伝達能力という観点から「語彙密度」(lexical density) (Halliday, 1989) に注目して、先行研究を概観する。さらに、実証的な先行研究から導き出された知見をもとに、予備的な調査として、即興性を伴うモノログ型タスクによる大学生英語学習者の発話データを分析する。

2. 先行研究

2.1. スピーキングにおける複雑さ、正確さ、流暢さ

L2 話者のスピーキング能力の査定は、1990 年代以降、タスクを用いた場合のインタラクション (task-based interaction) の分野で盛んに研究されている。タスクのタイプや難易度が L2 話者の発話にどう影響するかという問題は、スピーキングテストにおける能力判定基準とテスト課題の妥当性の検証 (e.g., Bachman, 1990) および TBLT (Task-based language teaching, タスクに基づく言語教育) (e.g., Ellis 2003; Nunan 2004; Robinson 2011) の 2 つの分野で詳しい分析が行われている。

L2 話者のスピーキング能力の発達を評価する指標として論じられるものに、「複雑さ」(complexity)、「正確さ」(accuracy)「流暢さ」(fluency)の 3 項目がある。この 3 項目は、しばしばそれぞれの語頭を取って CAF という用語でまとめて論じられる。CAF はスピーキングだけではなく、ライティング能力の評価にも用いられる場合があるが、本稿では、スピーキングに焦点を当てて先行文献を概括する。

Accuracy に重点を置いた Ellis (1987) の研究では、一連の漫画イラストを見て口頭説明をする課題で、事前準備 (pre-task planning) の機会が与えられるか否かで発話の正確さ

に差異が出たことが報告されている。一方、Crookes (1989)では、Ellis の研究で用いられたライティング活動を伴う準備活動を排し、準備の時間 (planning time) のみに注目した調査を行った。この Crookes の研究では、準備時間の有無で発話の正確さには変化は認められなかった一方、複雑さと流暢さで有意差が認められたとしている。

これら二つの研究を含めて検証し、3つの指標 (CAF) を整理して論じたのが、Skehan & Foster (1992) である。さらに Skehan (1998)、Skehan & Foster (1999)などの一連の論文で CAF の概念が繰り返し定義されているが、Skehan & Foster (1999)を例にとると、この3つの指標は表1の通り説明されている。

表1 Skehan & Foster (1999)における3つの指標の定義

指標	定義・説明
流暢さ (Fluency)	The capacity to use language in real time, to emphasize meanings, possibly drawing on more lexicalized systems 言語を即座に使用する能力。意味の伝達に重点が置かれ、語彙化された体系 (複数の形態素から一つの意義を持つ複合語、慣用句などの体系) を重用する傾向。
正確さ (Accuracy)	The ability to avoid errors in performance, possibly reflecting higher levels of control in the language 発話のなかで誤りを避ける能力。言語規則に沿ってことばを統御する能力を有していることが反映される。
複雑さ (Complexity)	The capacity to use more advanced language, with the possibility that such language may not be controlled so effectively 必ずしも高度に使いこなせていなくても、(語彙や統語などの側面において) より発展的な (複雑な) ことばを使用する能力。

Adapted from Skehan and Foster (1999), p. 93-120.

先行研究 (Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999) では、上の3つの指標は、相互に影響しあう関係にあることが論じられている。例えば、発話者が「正確さ」に注意すると、使い慣れていない語や構文の使用を避けるため、「複雑さ」に影響し、ポーズが増えて「流暢さ」にも影響を及ぼすという議論である。

日本でも、この3項目に基づいて中学生や高校生の英語スピーキング能力の発達に関する実証研究が行われており、概して Skehan (1998)や Skehan & Foster (1999)で論じられ

ていることが再現されている（小泉・栗寄 2002; 瀧口 2004; 小泉・山内 2003; 藤森 2004）。それらの実証研究で確認されているのは、具体的には、以下の点である。

- ・ タスク前のプランニングが、3つの指標（complexity、accuracy、fluency）で評価される発話の各側面に変化をもたらすこと。
- ・ 3つの指標はそれぞれ独立したものではなく、互いに影響しあう関係であること。
- ・ 発話者の「注意力資源」（attentional resources）がどう配分されるかによって、3指標の間で「トレード・オフ効果」（trade-off effects、一方を追求すると他方を犠牲にせざるを得ない現象）が生じること。

この「注意力資源」の有限性については、Robinson (2001) の認知仮説（Cognition Hypothesis）では、一部否定されている。Robinson によれば、難しいタスクの場合、complexity と accuracy にトレード・オフ効果は現れないとしている。ただし、fluency は下がるという。一方、Skehan & Foster (2012)の反論では、accuracy と complexity がトレード・オフせず共に高まる要因は、タスク・コンディションなどタスクの難易度以外の要因から影響を受けるとしている。このように、トレード・オフ効果の現れ方については議論が分かれるが（Housen & Kuiken, 2009, p.10; Skehan 2009, p.512; Skehan & Foster 2012）、トレード・オフ効果が fluency とそのほか二つの指標との間で何らかの形で存在するということは、これまでの調査研究から確認されており、語学学習者や教師の経験的知識とも矛盾しないと思われる。

2.2 語彙密度 (Lexical Density)

スピーキング能力の評価については、上述の CAF に「語彙」(lexis) を加え、4つの指標とする（Bachman & Palmer, 1996; Skehan & Foster 2012）のが近年の研究では一般的である。一定量のテキストで意味を伝達する能力の指標に語彙密度があるが、語彙密度はコーパス言語学、自然言語処理の分野と深く関わる複雑な概念である。紙幅の都合で、本研究のデータ分析に関係する基本的な情報のみを以下にまとめる。

語彙密度は、話す、または書くという形で算出された言語のパフォーマンスを計量的に測定する指標である。スピーキング、あるいはライティングで産出されたすべての語のうち、名詞、動詞、形容詞、副詞など内容語 (lexical words または content words) が占める割合を指す。密度が高いほど、一定量のテキスト内での意味の伝達量が大きいことを示す。語彙密度 (Lexical density) を算出するには、分析しようとするテキスト内に現れる名詞や動詞、形容詞などの内容語 (content words/lexical words) をすべての語の数で割り、100 をかけてパーセンテージで示す。式で表すと次のようになる。

$$\text{Lexical Density} = \frac{\text{The number of lexical words within the analyzed text}}{\text{The total number of words in the analyzed text}} \times 100$$

延べ語数に対し異なり語数（ある単語が、テキスト内で複数回使われている場合も 1 つとして数える）がどれほどあるかを示す Type/Token Ratio (TTR) は、語彙の豊富さを示す指標で、以下のような単純な数式で表される。

$$TTR = \frac{V}{N} \quad N=\text{総語数、} V=\text{異なり語数}$$

上の数式は、産出されたすべての語のうち、どれだけ異なった語が使われているかということを示す。使われている語の種類が多いと数値が大きくなり、語彙の豊富さが示され、数値が小さいと、繰り返して同じ語が使われていることになり、語彙が乏しいということの意味する。ただし、人間が算出する多様なことばを一つの計算式で分析するのは、単純すぎるとして批判も多い。

TTR を語彙の豊富さや効率的に意味を伝える能力の指標として用いるには大きく分けて二つの問題がある。まず、何をもって 1 token とするかは、かなり複雑な判断を必要とする場合がある。たとえば、名詞と名詞を重ねて 1 つの名詞のように使われている語を、複合名詞として 1 token と数えるか、前の名詞を形容詞的に使っていると考えて、後ろの名詞とあわせて 2 tokens とするかは、形態素解析をしても判定が困難な場合がある。

さらに大きな問題は、Baayen (2001) や他の多くの研究者が指摘するように、異なり語数は延べ語数と比例して増えることはなく、テキストが短い方が TTR は高い数値となる傾向があるということである。つまり、TTR はサンプルのサイズに影響を受けるので、長さが異なる文書・発話を比較するには問題があるとされている。

第二言語習得の分野では、TTR は L2 学習者の作文から語彙の豊富さを分析する手法として用いられたが、Wolfe-Quintero, et al. (1998) もサンプルの長さで値が左右されることを指摘し、延べ語数を調整して計算を行う改良型指標 (Carroll, 1967) の使用を提案した。しかし、改良型であっても延べ語数の異なる発話や文章を一律に同じ数式で計算するのは妥当ではない (Malvern & Richards, 2002) との批判が依然としてある。これまで、Carroll の指標のほか、コーパス言語学や自然言語処理の分野で、TTR を改良した様々な指標や、テキストの長さが語彙密度に与える影響を調整するパラメータ（例えば、D として知られる Malvern & Richards 2002 による補助変数）が何種類か提案されているが、いまだ多様なテキストに対応する指標についての統一的な理論はない。

一方、モノログ型タスクを用いて一定の時間制限のなかで採取された発話であれば、少なくとも時間的にはサンプルサイズが統一される。本研究では、学習者の発話能力の発

達を評価するための予備的な調査として、自分の家族について 45 秒間話す、というモノログ型タスクを用いて、「pruned token」（発話に現れた全語彙から繰り返しや言い直しを除いたもの）を中心に、語彙の使用に表れる話者の流暢さについてデータの分析を行う。

3. 調査の概要

3.1. 目的

今回の調査では、モノログ型タスクを使用し、日本人大学生の英語のスピーキング能力が 3 ヶ月で変化するかどうかを確認した。L2 話者のスピーキング能力の評価に広く使われている指標（CAF および語彙）のうち、本研究では意味伝達の効率という視点から流暢さを捉え、語彙の現れ方に注目する。授業の一環で採取する音声データを使用するため、変数を統制した実験ではなく、観察となる。

データ分析に当たっては、先行研究で用いられている分析観点を参考に、次のようにリサーチ・クエスチョンを立てた。

1. 時間制限のあるモノログ型タスクで算出される L2 発話で、算出される単語の数と意味の伝達量において 3 ヶ月間で観察可能な変化が起こるか。
2. 参加者が使用する語彙の多様性（異なり語数）は 3 ヶ月間で変化が認められるか。

3.2. データ収集の機会と参加者

本研究で用いるデータは、大学の半期（15 週）完結の英語選択科目（Special Topics: Pronunciation Workshop）の授業で録音した音声である。授業の形式や内容に関する情報は、三谷・辻田（2012; 2013）と三谷（2017）で詳しく報告している。教師から学習者へのフィードバックと学習者の変化の記録を目的として、学期の 3 週目と 14 週目に行っている発音プロファイルの音声の一部を、今回の分析に用いる。

データを収集した授業は選択科目のため、1 年生から 4 年生まで、様々な英語習熟度の学生が履修している。今回使用するデータは、ドイツ語またはフランス語を専門とする外国語学部の 1 年生で、3 週目（4 月）と 14 週目（7 月）の 2 回の録音に参加し、かつ研究に参加する意思を同意書で示した 18 人から収集したものである。

3.3. データ収集の手順

録音には CAL (Computer-assisted learning) 教室の機材を用いた。録音開始および停止の機械操作は教師が行い、学習者は各自マイク付きヘッドセットを装用し、スタートの合図とともに一斉に話し始めるよう指示されている。発音プロファイルは、Part A から Part C までの 3 つのパートから成り、Part B と C はモノログ型タスクである（表 2）。

表 2 発音プロフィール手順

パート	内容	時間制限	録音前準備・手順
Part A	Reading a paragraph (パラグラフ読み上げ)	なし	録音前に全体に目を通すために 8 分程度の準備時間を与える。
Part B	B-1 Talk briefly about your family. (家族について)	45 秒	B1 の録音開始前に、3 つのトピックについてまとめて事前準備に 5 分を与える。メモはいいが、「文」を予め作って書いておくことは禁止。B1 と
	B-2 Describe your school or work background and your purpose for studying English. (学校や仕事、英語を学習する目的について)	45 秒	B2、B2 と B3 の間は、録音された音声を一齐回収し保存する作業のため約 1 分の間がある。
	B-3 Describe your most pleasant childhood memory. (子供時代の最も楽しかった思い出について)	45 秒	
Part C	Introducing your partner (パートナー紹介)	45 秒	ペアでお互いに質問し合いながら、相手の情報を収集したあと、モノログ形式で録音。メモを取ってもいいが「文」を事前に作らないよう指示。

Adapted from Grant (2017), pp.3-4.

本稿で分析に用いるデータは、上の B-1 (Talk briefly about your family) の録音データである。Part B には 3 つのトピックがあり、各トピックについて 45 秒間、即興的に話し続ける。モノログ型タスクは自然な会話ではないが、話し相手の対応に影響されず、長めの発話を引き出すことができることから、流暢さの評価に適していると考えられる。参加者に与えられる事前準備の時間は、Part B の 3 つのトピック (表 2 参照) について何を話すかを考えるための 5 分程度である。何について話すか考えてメモしてもいいが、予め文を作っているといけない、と指示している。

3 週目と 14 週目のタスクは同一である。同じトピックを 2 度用いると、2 回目が 1 回目の練習のようになりかねず、2 回目のデータに影響が出る可能性を完全に否定することはできない。しかし、今回の分析に用いるトピックが特殊な知識を必要としない一般的なもの (自分の家族) であること、1 回目と 2 回目の間が約 3 ヶ月空いていることから、1 回目

に話したことが 2 回目に影響する可能性は低いと判断している。また、トピックが異なれば、話しやすさや必要とされる語彙に差が出るが、同じトピックではそのような差が生じない。いずれにせよ、自然発生的な発話データの比較においては、変数を完全にコントロールしてデータを取ることは容易ではなく、研究上の制限を開示した上で調査結果を示し、分析を行う必要があると考える。

3.4. データ分析

3.4.1. 分析の指標

限られた時間内で意味を効率よく伝達することを目指すモノログ型タスクの特性と、教師が使用できる簡易かつ有効な査定方法の作成という本研究の最終目標を念頭に、評価の観点を絞って分析を行った。小泉&栗寄（2002）で使用された 4 指標 10 項目の分析観点のうち流暢さに関わる項目、Lennon（1990）と Ortega（1999）による流暢さの測定項目を参考に、表 3 の通り分析の観点を定めた。

表 3 評価観点と客観的指標

評価観点	客観的指標
語彙	(1) Unpruned token の数（総語数）
	(2) Pruned token の数 = 総語数 : unpruned token 数から直後の繰り返し・言い直しを除いた数
	(3) 内容語の数
	(4) 内容語の Type の数（異なり語数） = 同じ語が複数回出現しても 1 と数える。語彙の豊富さを表す。
流暢さ	(5) Pruned speech rate = Pruned token の数 ÷ 話した秒数 × 60 繰り返しや自己訂正を除き 1 分間に換算した場合の語数
	(6) Lexical density (語彙密度) = 内容語の数 ÷ unpruned token の数 × 100

本研究では、「流暢さ」の指標のうち、まず pruned token の数の変化を起点として、学習者の発話を分析する。pruned token に注目するということは、一定の時間内でどれだけ意味のある情報を伝達しうるか、ということであり、本研究で用いた課題（与えられたトピックについて 45 秒間話し続ける）の評価に適していると考えられる。

3.4.2. データ分析の手順

B-1 のタスクの録音時間は 45 秒だが、録音開始からでなく最初に声が発せられてからの

40 秒間を分析に用いることとし、発話の書き起こしを行った。一斉スタートの合図を受けてから、いったん呼吸をしたり、マイクの位置を調整したりしてから話し始める参加者が多く、最初の 1～3 秒は周囲の雑音だけが録音されている事例が少なからずあったためである。

判定については、本研究者とは別の大学英語教員に依頼し、音声データを聞きながら書き起こしに間違いや抜けた語がないかどうかをチェックしてもらい、指摘のあった箇所については再度本研究者が確認を行って必要な修正を行った。

4. 結果

上述の評価観点と客観的指標（表 3）をもとにデータの分析を行った結果を、下の表 4 にまとめる。1 回目（3 週目）と 2 回目（14 週目）で、(1) の総語数（unpruned token）を見ると、2 回目の方が増加している（38.21→40.68）が、(4) の異なり語数で 2 回目は微減（15.47→15.42）となっている。これは、2 回目の方が発する語数は増えているが、使用語彙の多様性に変化はないことを示すので、平均値で見ると、語彙の面でのスピーキング能力の変化は確認できない。

一方、流暢さの指標のうち (5) の言い直しや繰り返しを除いて算出する「1 分あたりの語数」（pruned token÷話した秒数×60）では、1 回目 53.45 語、2 回目 57.00 語で、平均して 3.5 語ほど増えている。しかし、(3) の内容語の数を見ると、1 回目 19.89 語、19.47 語でほぼ同じであり、(6) 総語数に対する内容語の割合（語彙密度）を計算すると、1 回目 52.2%、2 回目 47.7%でわずかに低下している。今回の分析によると、意味伝達の効率は 3 ヶ月で向上していないという結果となる。

表 4 1 回目（第 3 週）と 2 回目（第 14 週）の発話の比較

評価観点		客観的指標	1 回目		2 回目	
			Mean	SD	Mean	SD
語彙	(1) Unpruned token の数		38.21	11.71	40.68	7.53
	(2) Pruned token の数		35.65	12.01	38.00	8.56
	(3) 内容語の数		19.89	6.67	19.47	4.98
	(4) 内容語の Type の数 (異なり語数)		15.47	4.48	15.42	3.86
流暢さ	(5) Pruned speech rate (1 分間に換算)		53.45	18.01	57.00	12.85
	(6) Lexical density (%)		52.2	8.41	47.7	6.97

さらに、限られた時間内でどれだけの情報を伝えられるかを詳しく見るため、表 4 で示

された数値を使って、内容語の現れ方を分析する。**pruned token** の数に対して内容語が占める割合、また使用されている内容語の異なり語数を算出したところ、次のような結果となった。

表 5 1 回目 (第 3 週) と 2 回目 (第 14 週) 内容語の割合と異なり度の比較

評価項目	1 回目	2 回目
(7) Pruned token に対する内容語の割合 (内容語の数÷Pruned Token の数)	56.2%	51.5%
(8) 内容語の中で同じ語が繰り返されない割合 (内容語の Type 数÷内容語の総数)	79.5%	80.3%

表 6 1 回目 (第 3 週) と 2 回目 (第 14 週) 1 分間に換算した場合の内容語数の比較

評価項目	1 回目		2 回目	
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
(9) 1 分間に換算した場合の内容語数 (内容語÷話した秒数×60)	29.84	10.0	29.21	7.47
(10) 1 分間に換算した場合の異なり内容語数 (内容語の Type÷話した秒数×60)	23.21	6.72	23.13	5.79

表 5 の通り、(7) の言い直しや直後の繰り返しを除いたすべての語 (**pruned token**) の中で内容語が占める割合は、1 回目が 56.2%、2 回目が 51.5%であったが、表 4 の (3) の内容語の数を見ると 2 回ともほぼ同じ (19.89 語と 19.47 語) なので、内容語による意味の伝達量は 1 回目、2 回目で差がないといえる。(7)は、(6) と計算式は似ているが、(6) が **unpruned token** の総数に対しての内容語の割合を示すのに対し、(7) は **pruned token** に対しての割合である。即興性のあるタスクであるため、同じ内容語を何度も繰り返して言い直す場合があるが (例: 「My father like, like, likes fish, fishing.」、(6) の計算式では伝達する意味の量は言い直しのない場合 (例: My father likes fishing.) と変わらないのに数値が上がってしまう。一方、繰り返しなどを省いて意味の伝達量を見ることを目指しているのが (7) である。

(8) は、言い直しや直後の繰り返しを除いた語の中から内容語のみを数え、そのなかで同じ内容語が繰り返されない割合を計算した。この数値が大きいと異なり度合いが低く、話者の語彙の多様性の高さ、つまり語彙の豊富さを示唆し、低いと同じ語を繰り返し使用

している、つまり語彙が乏しいことを意味する。結果は、1 回目が 79.5%で 2 回目が 80.3%であり、3 ヶ月間で目立つ変化はなかった。

最後に、表 6 の (9) では、1 分間に換算するといくつの内容語が含まれることになるのか、(10) は 1 分間あたりいくつ異なった内容語を使った計算になるかを示す。平均で見ると (9)、(10) とも 1 回目と 2 回目ではほとんど変化がないが、標準偏差を見ると 1 回目より 2 回目の方が参加者間の差が縮小していることが読み取れる。個別の学習者の変化や、異なるタスクのトピックを用いた場合との比較などに有効な計算式であると考えられる。

5. 分析結果についての考察と今後の研究課題

データから得られた数値から導き出されたリサーチ・クエスションの答えは次の通りである。

1. 時間制限のあるモノログ型タスクで算出される L2 発話で、算出される単語の数と意味の伝達量において 3 ヶ月間で観察可能な変化が起こるか。

総語数は、unpruned、pruned とともに、2 回目の方がわずかながら増えている（表 4 の(1)・(2)）。意味の伝達量を内容語の数や内容語の異なり語数で見ると、2 回目より 1 回目の平均値の数値がわずかに大きい、ほとんど差がない（表 4 の (3)・(4)）。

2. 参加者が使用する語彙の多様性（異なり語数）は 3 ヶ月間で変化が認められるか。

モノログ型タスクの中で使われた内容語の種類（異なり語数）を見ると（表 4 の(4)、表 6 の (10)）、1 回目と 2 回目ではほぼ変化がなかった。

1 回目に比べて 2 回目に総語数が増えたが、内容語が増えなかった理由として、2 回目は 1 回目より文法を整える機能語をうまく使っている可能性が考えられる。しかし、今回の分析では文法的な正確さを分析していないので、今後の研究における課題としたい。

教えている間は意識しなかったが、データ分析を通して気づいたことが 2 点ある。1 点目は、標準偏差の数値（表 4、表 6）が 1 回目に比べて 2 回目ですべて低くなっていることである。リサーチ・クエスションの答えは、「3 ヶ月でスピーキング能力における発達は参加者の平均から観察できない」というものであったが、産出する総語数や内容語の数につき、2 回目は参加者間でばらつきが少なくなっている。

2 点目は、一般に言われている話し言葉の平均的な語彙密度（40%未満）より、今回のデ

ータ分析で得た数値がかなり高いことである（1 回目 52.2%、2 回目 47.7%）。書き言葉に比べ、話し言葉は語彙密度が低いと言われているが（Halliday, 1989）、参加者が話した内容をあらためてオリジナルのデータで確かめると、この数値の高さはタスクのトピックに深く関係しているのではないかと推測できる。発話の中で、*father*, *mother* などの名詞が代名詞 *he*, *she* とすべきところで繰り返し使用されるため、発話全体に占める内容語の割合が話し言葉にしては高くなる一因となっている。この人称代名詞の使用頻度の低さは、日本語で家族を「彼」「彼女」と呼ぶことがないことが影響している可能性があり、人称代名詞の使用は、自然な英語という観点から、今後指導していくべき事項の一つであることを示唆している。

本研究は、L2 発話における「流暢さ」の概念に基づいて、使用語彙の数と種類に注目し、モノログ型タスクを用いた発話データの分析を試みたものである。教師が授業を行いつつ、入手可能なデータを分析し、自己査定を行う実践的なツール作りを目指すという観点から、まず第一歩として語彙のみに注目した。しかし、スピーキング能力の査定には、CAF に代表される指標の他、ポーズの現れ方（英語母語話者に特有の *mid-clause* ではなく、非母語話者に特徴的な *end-of-clause (or end-of-AS unit) pausing* がどのくらい見られるか）など、流暢さに影響する要素が数多くある。各指標の妥当性を先行研究に照らしながら検証し、総合的かつ実践的な評価方法の確立をめざすこと、また、タスクの種類や要求される難易度などの影響を考慮しつつ、いかに客観的な査定法を確定するかが今後の課題となる。

また、今回行ったのは授業内で行う小規模な調査で、統計的な処理に適するサンプルサイズではなかった。小規模な調査でいかに有意性の判定を行うかについても、検討を続ける必要がある。

さらに、今回の分析では、参加者全体の平均で見るとはっきりとした変化は見られなかったものの、1 回目と比較して 2 回目はデータの散らばり具合が縮小していることが観察された。今後、個別の参加者の発話の質的な変化に注目して、データの分析を深めていきたいと考えている。

参考文献

- Baayen, R. H. (2001). *Word frequency distributions* (Vol. 18). Dordrecht: Springer Science & Business Media.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford university press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and*

- developing useful language tests* (Vol. 1). Oxford: Oxford University Press.
- Crookes, G. (1989). Planning and interlanguage variation. *Studies in second language acquisition*, 11(4), 367-383.
- Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in second language acquisition*, 9(1), 1-19.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied linguistics*, 30(4), 474-509.
- Ellis, R., & Barkhuizen, G. P. (2005). *Analysing learner language*. Oxford: Oxford University Press.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second language acquisition*, 18(3), 299-323.
- 藤森千尋 (2004). 「スピーチプロダクションの測定方法: 正確さ, 流暢さ, 複雑さ」『関東甲信越英語教育学会研究紀要』, 18, 41-52.
- Grant, L. (2017). *Well said: Pronunciation for clear communication*. Boston, MA: Thomson.
- Halliday, M. A. K. (1989). *Spoken and Written Language (Language Education)*. Oxford: Oxford University Press.
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied linguistics*, 30(4), 461-473.
- 小泉利恵 & 栗寄逸美 (2002). 「日本人中学生のモノログにおけるスピーキングの特徴」. 『関東甲信越英語教育学会研究紀要』, 16, 17-28.
- 小泉利恵 & 山内逸美 (2003). 「日本人中学生のスピーキング能力の発達: 自己紹介のタスクを用いて」. 『関東甲信越英語教育学会研究紀要』, 17, 33-44.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language learning*, 40(3), 387-417.
- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language testing*, 19(1), 85-104.
- Mann, S., & Walsh, S. (2017). *Reflective practice in English language teaching: Research-based principles and practices*. New York: Routledge.
- McDonough, J., & McDonough, S. (2014). *Research methods for English language*

teachers. Routledge.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20, 83-108.

三谷裕美 (2017). 「英語発音指導が日本人学習者のリズムの習得に与える影響: 大学における発音授業の事例研究」. 『獨協大学外国語教育研究所紀要』, 6, 41-58.

三谷裕美 & 辻田麻里 (2012). 「大学における英語発音指導: ニーズ分析と授業実践報告」. 『獨協大学外国語教育研究所紀要』, 1, 137-149.

三谷裕美 & 辻田麻里 (2013). 「中間言語音韻の分節素性と超分節素性: 日本人大学生の英語発音の事例研究」 獨協大学外国語教育研究所紀要』, 2, 39-62.

Nunan, D. (2004). *Task-based language teaching: A comprehensive revised edition of designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in second language acquisition*, 21(1), 109-148.

Richards, J. C., & Farrell, T. S. C. (2005). *Professional development for language teachers: Strategies for teacher learning*. New York: Cambridge University Press.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied linguistics*, 22(1), 27-57.

Robinson, P. (2011). *Task-based language learning* (Vol. 61). Malden, MA: John Wiley & Sons.

Schon, D. A. (1984). *The Reflective Practitioner: How professionals think in action*. New York: Basic Books.

Skehan, P. (1998). Task-based instruction. *Annual review of applied linguistics*, 18, 268-286.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied linguistics*, 30(4), 510-532.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language teaching research*, 1(3), 185-211.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language learning*, 49(1), 93-120.

Skehan, P., & Foster, P. (2012). Complexity, accuracy, fluency and lexis in task-based performance. *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, 32, 199.

- 瀧口均 (2004). 「日本人 EFL 中学生のスピーキング能力の発達研究:『流暢さ』『複雑さ』『正確さ』の指標を用いて」. 『関東甲信越英語教育学会研究紀要』, 18, 1-13.
- Wolfe-Quintero, K., Inagaki, S., Kim, H. Y., Kim, H. Y., & Inagaki, S. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity* (p. 187). Honolulu: University of Hawai'i Press.